



Johnson, O. (2017). A discrete log-Sobolev inequality under a Bakry-Émery type condition. *Annales de l'Institut Henri Poincaré (B) Probabilités et Statistiques*, 53(4), 1952-1970.
<https://doi.org/10.1214/16-AIHP778>

Publisher's PDF, also known as Version of record

Link to published version (if available):
[10.1214/16-AIHP778](https://doi.org/10.1214/16-AIHP778)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the final published version of the article (version of record). It first appeared online via Project Euclid at <https://projecteuclid.org/euclid.aihp/1511773733> . Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

A discrete log-Sobolev inequality under a Bakry–Émery type condition

Oliver Johnson

School of Mathematics, University of Bristol, University Walk, Bristol, BS8 1TW, UK. E-mail: maotj@bristol.ac.uk

Received 8 July 2015; revised 4 July 2016; accepted 6 July 2016

Abstract. We consider probability mass functions V supported on the positive integers using arguments introduced by Caputo, Dai Pra and Posta, based on a Bakry–Émery condition for a Markov birth and death operator with invariant measure V . Under this condition, we prove a new modified logarithmic Sobolev inequality, generalizing and strengthening results of Wu, Bobkov and Ledoux, and Caputo, Dai Pra and Posta. We show how this inequality implies results including concentration of measure and hypercontractivity, and discuss how it may extend to higher dimensions.

Résumé. Nous considérons des distributions de probabilité V à support dans l'ensemble des entiers positifs, en utilisant des arguments introduits par Caputo, Dai Pra et Posta, basés sur une condition de Bakry–Émery pour une chaîne de naissance et mort avec mesure invariante V . Sous cette condition, nous prouvons une nouvelle inégalité de Sobolev logarithmique modifiée, en généralisant et améliorant des résultats de Wu, de Bobkov et Ledoux, et de Caputo, Dai Pra et Posta. Nous montrons comment cette inégalité implique des résultats tels que la concentration de la mesure et l'hypercontractivité, et nous discutons d'une extension possible aux dimensions supérieures.

MSC: 39B62; 60J10

Keywords: Bakry–Émery condition; Birth and death chain; Concentration of measure; Discrete probability measure; Log-concavity; Log-Sobolev inequality

1. Introduction and main results

In their classic 1985 paper [2], Bakry and Émery introduced the Γ_2 operator and showed that (assuming the Bakry–Émery condition) it could be used to prove results such as Poincaré and log-Sobolev inequalities. The books by Bakry, Gentil and Ledoux [3] and by Guionnet and Zegarlinski [16] review this theory. We give a brief overview in Section 2.

Historically, this analysis was restricted to functions on continuous spaces such as \mathbb{R}^d , or more generally Riemannian manifolds. However, more recently it was extended to discrete spaces by Caputo, Dai Pra and Posta [7], by considering the evolution of a birth and death Markov chain. This analysis built on the work of Chen and others (see e.g. [10]), who used properties of birth and death chains to bound the spectral gap. We use a version of these methods here; we fix probability mass function V whose support is the whole of \mathbb{Z}_+ and fix the generator L_V of a birth and death Markov chain with invariant measure V :

Definition 1.1. Given a fixed probability mass function V , write L_V for the operator acting on functions, and L_V^* for the adjoint with respect to counting measure by:

$$L_V f(x) := (f(x+1) - f(x)) - \frac{V(x-1)}{V(x)}(f(x) - f(x-1)) \quad (1)$$

and

$$L_V^* f(x) := f(x-1) - \left(1 + \frac{V(x-1)}{V(x)}\right) f(x) + \frac{V(x)}{V(x+1)} f(x+1). \quad (2)$$

In Equation (1) and throughout, we use the convention that $V(-1) = 0$.

In Section 3 we formally define the resulting operators $\Gamma_1^{(V)}$ and $\Gamma_2^{(V)}$ and discuss the resulting Bakry–Émery condition (we refer to this as the integrated BE(c) condition, Condition 3). In contrast to the continuous case we require average control, rather than pointwise control of the relevant operators. However (see Remark 4.3), this $\Gamma_1^{(V)}$ operator does not satisfy a product rule, so classical proofs of log-Sobolev inequalities do not carry over.

The continuous Bakry–Émery condition relates to log-concavity of the underlying reference measure, with the Gaussian playing a distinguished role (see Example 2.2). We use a similar condition here, which corresponds to Assumption A of [7] (in the case where the upward jump rates of the birth and death chain are all equal to 1):

Definition 1.2. Given a probability mass function V , write

$$\mathcal{E}^{(V)}(x) := \frac{V(x)^2 - V(x-1)V(x+1)}{V(x)V(x+1)} = \frac{V(x)}{V(x+1)} - \frac{V(x-1)}{V(x)}. \quad (3)$$

Condition 1 (c -log-concavity). If $\mathcal{E}^{(V)}(x) \geq c$ for all $x \in \mathbb{Z}_+$, we say that V is c -log-concave.

In Proposition 4.2 we show the integrated BE(c) condition is implied by c -log-concavity. [7] showed that c -log-concavity follows from the ultra log-concavity (ULC) property of Liggett and Pemantle [25,28]. Hence integrated BE(c) holds for parametric families of random variables including Poisson, binomial and Poisson-binomial (Bernoulli sums). For the Poisson mass function with mean λ (we write $V = \Pi_\lambda$), the $\mathcal{E}^{(V)}(x) \equiv 1/\lambda$, so c -log-concavity holds with $c = 1/\lambda$, which Lemma 5.3 shows is an extreme value. This helps us to understand $\mathcal{E}^{(V)}(x)$ as a discrete curvature term, in the sense discussed by Chafaï in [8,9]. For example, in [8, Section 1.3] it is remarked that the $M/M/\infty$ queue (corresponding to $V = \Pi_\lambda$) can be understood to have constant curvature.

As in [2] and [7], in Section 6 we prove a new (modified) log-Sobolev inequality, Theorem 1.3, which is the main result of this paper. In the continuous case, the log-Sobolev inequality holds under the Bakry–Émery condition (see e.g. Theorem 2.5), whereas our result requires the (slightly stronger) c -log-concavity condition.

To fix notation, we write $\Delta f(x) = f(x+1) - f(x)$ for the right difference operator, and $Lf(x) = f(x+1) - 2f(x) + f(x-1)$. Given a probability mass function V and function f , we write $\text{Var}_V(f) = \sum_{x=0}^{\infty} V(x)(f(x) - \mu_{V,f})^2$, where $\mu_{V,f} = \sum_{x=0}^{\infty} V(x)f(x)$. Similarly, we write $\text{Ent}_V(f) = \sum_{x=0}^{\infty} V(x)f(x)\log f(x) - \mu_{V,f}\log \mu_{V,f}$.

We now state the main result of this paper. As discussed in more detail in Remark 1.4, the form of this inequality is suggested by the fact that it holds for the case where V is Poisson, as proved in [32, Theorem 1.1] and [8, Corollary 2.4].

Theorem 1.3 (New modified log-Sobolev inequality). Fix probability mass function V , whose support is the whole of the positive integers \mathbb{Z}_+ and which satisfies the c -log-concavity condition (Condition 1). For any function f with positive values:

$$\text{Ent}_V(f) \leq \frac{1}{c} \sum_{x=0}^{\infty} V(x)f(x+1) \left(\log \left(\frac{f(x+1)}{f(x)} \right) - 1 + \frac{f(x)}{f(x+1)} \right). \quad (4)$$

By the well-known bound $\log 1/u - 1 + u \geq 0$ for all $u > 0$, the right-hand side of (4) is positive.

Previous work on different forms of log-Sobolev inequalities in discrete settings is discussed and summarised by Bobkov and Tetali [5]. In particular, [5, Proposition 3.6] gives a hierarchy of different constants and discusses the implications between them. One particular form of interest is the more standard modified log-Sobolev inequality:

$$\text{Ent}_V(f) \leq C \sum_{x=0}^{\infty} V(x)(f(x+1) - f(x))(\log f(x+1) - \log f(x)), \quad (5)$$

also considered in [6] and discussed in Remark 1.4, item 7.

Remark 1.4. We discuss Theorem 1.3 in detail, to put it in context:

1. Suppose $f(x) = p(x)/V(x)$, for probability mass function p . Using normalizing constant $K = (\sum_{x=0}^{\infty} p(x+1)V(x)/V(x+1))^{-1}$, then $p^{\#}(x) = Kp(x+1)V(x)/V(x+1)$ is a probability mass function acting as a weighted version of p . Then (4) means relative entropy $D(p \parallel q) := \sum_{x=0}^{\infty} p(x) \log(p(x)/q(x))$ is bounded by the sum of two positive terms, as

$$D(p \parallel V) \leq \frac{1}{cK} \left(D(p^{\#} \parallel p) + \left(\log \frac{1}{K} - 1 + K \right) \right). \quad (6)$$

2. If $V = \Pi_{\lambda}$ and λ is the mean of p , then $p^{\#}$ is the size-biased version of p (see e.g. [12]), $c = 1/\lambda$ and $K = 1$. We recover the fact that

$$D(p \parallel V) \leq \lambda D(p^{\#} \parallel p), \quad (7)$$

which is a log-Sobolev inequality of Wu [32, Theorem 1.1], reproved more directly in [33] (see also [8, Corollary 2.4]). The relationship between Theorem 1.3 and (7) is the same as between the Bakry–Émery log-Sobolev inequality (Theorem 2.5) and the original result of Gross [15].

3. The right-hand side of (4) can be understood to be

$$\frac{1}{c} \sum_{x=0}^{\infty} V(x) A^{\Phi}(f(x), f(x+1) - f(x)), \quad (8)$$

where $A^{\Phi}(u, v) = \Phi(u+v) - \Phi(u) - \Phi'(u)v$ is the A -transform of the function $\Phi(u) := u \log u$, as introduced by Chafaï in [8]. This allows us to understand the fact that Theorem 1.3 reduces to [8, Corollary 2.4] in the case where $V = \Pi_{\lambda}$.

4. Using the bound $\log(f(x+1)/f(x)) \leq f(x+1)/f(x) - 1$ we deduce that if V satisfies the c -log-concavity condition then:

$$\text{Ent}_V(f) \leq \frac{1}{c} \sum_{x=0}^{\infty} V(x) \frac{(f(x+1) - f(x))^2}{f(x)}. \quad (9)$$

Hence taking $V = \Pi_{\lambda}$ and $c = 1/\lambda$ we see Theorem 1.3 generalizes and strengthens the log-Sobolev inequality of Bobkov and Ledoux [4, Corollary 4].

5. In the spirit of [24] the right-hand side of (9) is a form of scaled Fisher information, equalling

$$\frac{1}{c} \sum_{x=0}^{\infty} p(x) \left(\frac{p(x+1)V(x)}{p(x)V(x+1)} - 1 \right)^2 = \frac{1}{c} \sum_{x=0}^{\infty} p(x) \left(\frac{p^{\#}(x)}{Kp(x)} - 1 \right)^2,$$

where we may interpret the ratio $p^{\#}(x)/(Kp(x))$ as a scaled score function.

6. Theorem 1.3 is sharp; equality is achieved in (4) when $V = \Pi_{\lambda}$ for any value of a in $f(x) = \exp(ax + b)$, or equivalently in (6) for $p = \Pi_{\mu}$. To verify this, note that $p(x)/V(x) = \exp(\lambda - \mu)(\mu/\lambda)^x$, so the left-hand side of (6) becomes $\lambda - \mu + \mu \log(\mu/\lambda)$. Further, $c = 1/\lambda$, $p^{\#} = p$ and $K = \lambda/\mu$, the right-hand side of (6) is $\mu(\log(\mu/\lambda) - 1 + \lambda/\mu)$ and equality holds.
7. Further, Theorem 1.3 strengthens the log-Sobolev inequality of Caputo et al. [7] who showed that (under the same condition) the modified log-Sobolev inequality in the sense of [5] and Equation (5) holds:

$$\text{Ent}_V(f) \leq \frac{1}{c} \sum_{x=0}^{\infty} V(x) (f(x+1) - f(x)) \log \left(\frac{f(x+1)}{f(x)} \right). \quad (10)$$

The expression (10) is a symmetrized version of (4), with its right-hand side equal to the right-hand side of (4) plus a similar-looking term (which is again positive, as before), namely

$$\frac{1}{c} \sum_{x=0}^{\infty} V(x) f(x) \left(\log \left(\frac{f(x)}{f(x+1)} \right) - 1 + \frac{f(x+1)}{f(x)} \right). \quad (11)$$

8. Again, to consider sharpness; [7] shows (10) is sharp, in the weaker sense that the constant cannot be improved in general. However, equality only holds in (10) for $f(x) = \exp(ax + b)$ in the limit as $a \rightarrow \infty$ (the term (11) vanishes in this limit).
9. In the case $V = \Pi_\lambda$, the right-hand side of (4) is strongly reminiscent of [4, Corollary 7], though that result is expressed in terms of the modulus of differences of functions and their logarithms, and is only sharp for $f(x) = \exp(ax + b)$ where $a \geq 0$.

In Section 8 we discuss some consequences of Theorem 1.3, including concentration of measure and hypercontractivity results. The assumption that V has support the whole of \mathbb{Z}_+ can be relaxed by a perturbation argument (see Remark 6.2). However, making this assumption simplifies the exposition of the paper.

In a standard way, we consider $f = (1 + \epsilon g)$, and let $\epsilon \rightarrow 0$ in (4). The left-hand side behaves like $(\epsilon^2/2) \text{Var}_V(g)$, and the right-hand side like $(\epsilon^2/2)(\frac{1}{c} \sum_{x=0}^{\infty} V(x) \Delta g(x)^2)$, so that as expected, the log-Sobolev inequality Theorem 1.3 implies a Poincaré inequality (12). This Poincaré inequality can be proven independently, and is in fact equivalent to the slightly weaker integrated BE(c) assumption, with the same constant (see Section 7).

Theorem 1.5 (Poincaré inequality). *Fix probability mass function V , whose support is the whole of the positive integers \mathbb{Z}_+ . Then for any function f :*

$$\text{Var}_V(f) \leq \frac{1}{c} \sum_{x=0}^{\infty} V(x) \Delta f(x)^2, \quad (12)$$

if and only if V satisfies the integrated BE(c) condition (Condition 3).

In the Poisson case where $V = \Pi_\lambda$, taking $c = 1/\lambda$ we recover the Poincaré inequality of Klaassen [22]. In general, as discussed in Section 7, Theorem 1.5 is comparable to a Poincaré inequality proved under similar conditions by very different methods in [12].

When V has finite support, we may adapt the Markov chain, and choose a different L_V with invariant distribution V . The correct statement of the Poincaré inequality in this context may be in the spirit of [18] (where we adapt the form of the derivative operator used). This remains a topic for future research.

We briefly describe some related work in the literature. The proof of Theorem 1.5 mirrors the type of argument given for a range of discrete systems, including a class of Markov dynamics that includes Kawasaki dynamics, by Boudou, Caputo, Dai Pra and Posta [6]. The work of Caputo, Dai Pra and Posta [7] was developed by Fathi and Maas [14], building on a Markov chain-based construction of Ricci curvature on a discrete space introduced by Erbar and Maas [13] (see also [27]). In particular, [13, Theorem 1.5] showed that Poincaré and modified log-Sobolev inequalities (in the form of (5)) hold assuming a bound on their form of Ricci curvature. A form of the c -log-concavity condition was used by Joulin [21] (along with another form of curvature condition), and was used to derive concentration of measure bounds in the context of birth and death processes.

2. Bakry–Émery calculus for continuous spaces

We briefly discuss the classical Bakry–Émery calculus for continuous spaces – see [3] for a clear and detailed review. Since Theorem 1.3 considers measures supported on \mathbb{Z}_+ , we restrict our description to measures on \mathbb{R} , although this theory holds in considerably greater generality. The key is a second-order differential operator L , self-adjoint with respect to reference measure $d\mu$, which allows the creation of the so-called carré du champ operator Γ_1 , and the iterated operator Γ_2 .

Definition 2.1. For any functions f and g , write

$$\Gamma_1(f, g) = \frac{1}{2} [L(fg) - fLg - gLf], \quad (13)$$

$$\Gamma_2(f, g) = \frac{1}{2} [L(\Gamma_1(f, g)) - \Gamma_1(f, Lg) - \Gamma_1(g, Lf)]. \quad (14)$$

The central definition in the theory is the following, which was introduced in [2].

Condition 2 (Bakry–Émery condition). We say that the Bakry–Émery condition holds with constant c if for all functions f :

$$\Gamma_2(f, f) \geq c\Gamma_1(f, f). \quad (15)$$

One key example is the following, which motivates the c -log-concavity property, Condition 1. It simplifies further if $U = cx^2/2$ and hence μ is Gaussian with variance $1/c$.

Example 2.2 (e.g. [16], Exercise 4.18). For function $U(x)$, take $d\mu(x) = \exp(-U(x)) dx/\mathcal{Z}$. Write

$$Lf(x) = f''(x) - U'(x)f'(x) = \exp(U(x)) (f'(x) \exp(-U(x)))'.$$

It is simple to verify that (for well-behaved U , including those satisfying $U''(x) \geq c$, as assumed below), L is self-adjoint with respect to μ , that $\Gamma_1(f, g) = f'g'$, and

$$\Gamma_2(f, g) = f''(x)g''(x) + U''(x)f'(x)g'(x). \quad (16)$$

If we assume that $U''(x) \geq c$ then $\Gamma_2(f, f) = f''(x)^2 + U''(x)f'(x)^2 \geq cf'(x)^2 = c\Gamma_1(f, f)$, and the Bakry–Émery condition 2 holds with constant c .

Remark 2.3. As discussed in [16], the Γ_1 operator satisfies a product rule of the form:

$$\Gamma_1(f, gh) = \Gamma_1(f, g)h + \Gamma_1(f, h)g. \quad (17)$$

As a result (see e.g. [16, Lemma 4.12]), for any well-behaved function v , the Γ_1 operator satisfies a chain rule of the form

$$\Gamma_1(v(f), g) = v'(f)\Gamma_1(f, g), \quad (18)$$

which is a key reason that the Bakry–Émery theory applies in the continuous case.

We state two results which arise from the Bakry–Émery calculus, as first described in [2] and reviewed and extended since by a variety of authors. For example, taking $U(t) = t^2$ in [2, Proposition 5] we deduce (see also [3, Proposition 4.8.1]):

Theorem 2.4. *If the Bakry–Émery condition (Condition 2) holds with constant c then the Poincaré inequality holds with constant $1/c$; that is for any function f ,*

$$\text{Var}_\mu(f) \leq \frac{1}{c} \int \Gamma_1(f, f)(x) d\mu(x).$$

Similarly [2, Theorem 1] (see also [3, Proposition 5.7.1]) gives that:

Theorem 2.5. *If the Bakry–Émery condition (Condition 2) holds with constant c then the logarithmic Sobolev inequality holds with constant $1/c$; that is for any function f with positive values:*

$$\text{Ent}_\mu(f) \leq \frac{1}{2c} \int \frac{\Gamma_1(f, f)(x)}{f(x)} d\mu(x).$$

If μ is Gaussian with variance σ^2 , since (as discussed in Example 2.2) we take $c = 1/\sigma^2$, and the right-hand side becomes the standardized Fisher information $\int f'(x)^2/f(x) d\mu(x)$, we recover the original log-Sobolev inequality of Gross [15] (see also Stam [30]).

3. Birth and death Markov chain

Fix a probability mass function V supported on the whole of \mathbb{Z}_+ . As in [7], we construct a birth and death Markov chain with invariant distribution V . In [7] more general upwards jump rates are considered, but this construction is sufficient for our purposes.

Definition 3.1. Define the birth and death Markov chain with upward jumps rate equal to 1, and downward jump rate at x equal to $V(x-1)/V(x)$. Equivalently, define the Q -matrix:

$$Q := \begin{pmatrix} -1 & 1 & 0 & 0 & \dots \\ \frac{V(0)}{V(1)} & -\frac{V(0)}{V(1)} - 1 & 1 & 0 & \dots \\ 0 & \frac{V(1)}{V(2)} & -\frac{V(1)}{V(2)} - 1 & 1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \quad (19)$$

We consider evolution of probability mass functions by $p_t := p \exp(tQ)$, so that for any x :

$$\frac{\partial}{\partial t} p_t(x) = p_t Q = p_t(x-1) - \left(1 + \frac{V(x-1)}{V(x)}\right) p_t(x) + \frac{V(x)}{V(x+1)} p_t(x+1) = L_V^* p_t(x), \quad (20)$$

using the notation of Definition 1.1.

Example 3.2. If V is Poisson Π_λ , then Equation (20) becomes

$$\frac{\partial}{\partial t} p_t(x) = p_t Q = p_t(x-1) - \left(1 + \frac{x}{\lambda}\right) p_t(x) + \frac{(x+1)}{\lambda} p_t(x+1),$$

as in [20, Equation (14)], giving the evolution of the $M/M/\infty$ queue. In [20], the action of this Markov chain was used to prove the maximum entropy property of the Poisson distribution, under the ultra-log-concavity condition (Condition 4).

Writing vector $\mathbf{V} = (V(0), V(1), V(2), \dots)$ the $\mathbf{V}Q = 0$, so V is indeed the invariant distribution of this Markov chain. Indeed, the Markov chain satisfies the detailed balance condition, and hence is reversible. Further, since V is supported on the whole of \mathbb{Z}_+ , the Markov chain is irreducible, and we deduce that this invariant measure is unique, meaning that the probabilities $p_t(x) \rightarrow V(x)$ as $t \rightarrow \infty$. Since the rate of upward jumps is constant, the chain is non-explosive, since the expected time to reach ∞ is $\sum_{x=0}^{\infty} 1/Q_{x;x+1} = \sum_{x=0}^{\infty} 1 = \infty$.

In fact, here it is more useful to consider the evolution of functions.

Definition 3.3. Given a function f , consider the sequence of functions f_t evolving as $\exp(tQ)f$, so that

$$\frac{\partial}{\partial t} f_t(x) = Qf_t(x) = f_t(x+1) - f_t(x) - \frac{V(x-1)}{V(x)}(f(x) - f(x-1)) = L_V f_t(x), \quad (21)$$

where L_V is the operator defined in Definition 1.1.

Next we give a result which allows us to prove the equivalent of Example 2.2.

Lemma 3.4. *Observe that for any functions f and g , rearrangement gives that L_V is self-adjoint with respect to V where, writing $\Delta f(x) = f(x+1) - f(x)$,*

$$\sum_{x=0}^{\infty} V(x) f(x) L_V g(x) = \sum_{x=0}^{\infty} V(x) L_V f(x) g(x) = - \sum_{x=0}^{\infty} V(x) \Delta f(x) \Delta g(x). \quad (22)$$

Proof. This follows by adjusting the index of summation since

$$\begin{aligned} \sum_{x=0}^{\infty} V(x) f(x) L_V g(x) &= \sum_{x=0}^{\infty} V(x) f(x) \left(g(x+1) - g(x) - \frac{V(x-1)}{V(x)} (g(x) - g(x-1)) \right) \\ &= \sum_{x=0}^{\infty} V(x) f(x) (g(x+1) - g(x)) - \sum_{x=0}^{\infty} V(x) f(x+1) (g(x+1) - g(x)), \end{aligned}$$

and the result follows. \square

4. Integrated Bakry–Émery condition

Given the operator L_V , we define the $\Gamma_1^{(V)}$ and $\Gamma_2^{(V)}$ operators induced by it in the standard way introduced by [2].

Definition 4.1. For any functions f and g , write

$$\Gamma_1^{(V)}(f, g) = \frac{1}{2} [L_V(fg) - fL_V g - gL_V f], \quad (23)$$

$$\Gamma_2^{(V)}(f, g) = \frac{1}{2} [L_V(\Gamma_1^{(V)}(f, g)) - \Gamma_1^{(V)}(f, L_V g) - \Gamma_1^{(V)}(g, L_V f)]. \quad (24)$$

We next introduce the Integrated Bakry–Émery condition; note that in contrast to the classical Bakry–Émery condition (Condition 2) we only require control of the average (with respect to V) of $\Gamma_2^{(V)}$ and $\Gamma_1^{(V)}$, not pointwise control.

Condition 3 (Integrated BE(c)). We say that probability mass function V satisfies the integrated BE(c) condition if for all functions f :

$$\sum_{x=0}^{\infty} V(x) \Gamma_2^{(V)}(f, f)(x) \geq c \sum_{x=0}^{\infty} V(x) \Gamma_1^{(V)}(f, f)(x). \quad (25)$$

Proposition 4.2. *For any f and g , writing $Lf(x) = f(x+1) - 2f(x) + f(x-1)$ we deduce:*

$$\sum_{x=0}^{\infty} V(x) \Gamma_1^{(V)}(f, g)(x) = \sum_{x=0}^{\infty} V(x) (f(x+1) - f(x)) (g(x+1) - g(x)), \quad (26)$$

$$\begin{aligned} \sum_{x=0}^{\infty} V(x) \Gamma_2^{(V)}(f, g)(x) &= \sum_{x=0}^{\infty} V(x) Lf(x+1) Lg(x+1) \\ &\quad + \sum_{x=0}^{\infty} V(x) \mathcal{E}^{(V)}(x) (f(x+1) - f(x)) (g(x+1) - g(x)). \end{aligned} \quad (27)$$

Hence, if V is c -log-concave (if $\mathcal{E}^{(V)}(x) \geq c$ for all x) then the integrated BE(c) condition holds.

Proof. Observe that, the $\mathcal{E}^{(V)}$ term naturally emerges here and defines a curvature term, since (28) expresses the difference between two adjacent derivatives:

$$L_V f(x+1) - L_V f(x) = Lf(x+1) - Lf(x) \frac{V(x-1)}{V(x)} - \mathcal{E}^{(V)}(x)(f(x+1) - f(x)). \quad (28)$$

Using Lemma 3.4, since $\sum_{x=0}^{\infty} V(x) L_V h(x) = 0$ for any function h , we know

$$\sum_{x=0}^{\infty} V(x) \Gamma_1^{(V)}(f, g)(x) = - \sum_{x=0}^{\infty} V(x) f(x) L_V g(x), \quad (29)$$

and (26) follows by (22). Multiplying by $V(x)$ and summing, we recover (26) (as suggested by Lemma 3.4). Using (26), similarly we know that $\sum_{x=0}^{\infty} V(x) \Gamma_2^{(V)}(f, g)(x)$ equals

$$-\frac{1}{2} \sum_{x=0}^{\infty} V(x) \Gamma_1^{(V)}(f, L_V g)(x) - \frac{1}{2} \sum_{x=0}^{\infty} V(x) \Gamma_1^{(V)}(L_V f, g)(x) \quad (30)$$

$$= - \sum_{x=0}^{\infty} V(x) (g(x+1) - g(x)) (L_V f(x+1) - L_V f(x)) \quad (31)$$

$$\begin{aligned} &= - \sum_{x=0}^{\infty} V(x) (g(x+1) - g(x)) Lf(x+1) + \sum_{x=0}^{\infty} V(x-1) (g(x+1) - g(x)) Lf(x) \\ &\quad + \sum_{x=0}^{\infty} V(x) \mathcal{E}^{(V)}(x) (g(x+1) - g(x)) (f(x+1) - f(x)) \\ &= \sum_{x=0}^{\infty} V(x) [Lf(x+1) Lg(x+1) + \mathcal{E}^{(V)}(x) (g(x+1) - g(x)) (f(x+1) - f(x))], \end{aligned} \quad (32)$$

where (31) follows by (26), since the two terms in (30) are both equal (as (29) shows that as usual, they can both be expressed as $\frac{1}{2} \sum_{x=0}^{\infty} V(x) L_V f(x) L_V g(x)$). The final result (32) follows on relabelling, having substituted (28) in the second term of (31). \square

Remark 4.3. Using (26) we deduce that $\Gamma_1^{(V)}$ only satisfies a modified form of the product rule in (17). That is since $g(x+1)h(x+1) - g(x)h(x) = h(x+1)(g(x+1) - g(x)) + g(x)(h(x+1) - h(x))$ we know that

$$\begin{aligned} \sum_{x=0}^{\infty} V(x) \Gamma_1^{(V)}(f, gh)(x) &= \sum_{x=0}^{\infty} V(x) (f(x+1) - f(x)) (g(x+1) - g(x)) h(x+1) \\ &\quad + \sum_{x=0}^{\infty} V(x) (f(x+1) - f(x)) (h(x+1) - h(x)) g(x). \end{aligned}$$

5. The c -log-concavity condition

The c -log-concavity property (Condition 1) corresponds to the bound $U''(x) \geq c$ discussed in Example 2.2. Condition 1 was introduced as Assumption A in [7], who showed that it is implied by the ultra-log-concavity condition of Pemantle [28] and Liggett [25]:

Condition 4 (ULC). If a probability mass function V has the property that V/Π_λ is a log-concave sequence, then we say that V is ultra-log-concave (ULC).

Lemma 5.1 ([7], Section 3.2). *If V is ULC, then it is c -log-concave, with $c = V(0)/V(1)$.*

Notice that if U and V are probability mass functions then

$$\frac{(U \star V)(1)}{(U \star V)(0)} = \frac{U(1)}{U(0)} + \frac{V(1)}{V(0)},$$

where $(U \star V)$ represents the convolution. In the light of Lemma 5.1 this suggests the conjecture that if U and V are c -log-concave with constants c_U and c_V respectively, then $(U \star V)$ is c -log-concave with constant $\geq (1/c_U + 1/c_V)^{-1}$. (Recall that Walkup [31] proved a result which implies that if U and V are ULC, then so is $(U \star V)$.)

We discuss probability mass functions V for which Condition 1 is satisfied. While Theorem 1.3 requires that V has support the whole of \mathbb{Z}_+ , it is still instructive to take V with finite interval support (see Remark 6.2).

Example 5.2.

1. If $V = \Pi_\lambda$ is Poisson, then since $V(x)/V(x+1) = (x+1)/\lambda$, we know that $\mathcal{E}^{(V)}(x) \equiv 1/\lambda$, so V is c -log-concave (with equality), with $c = 1/\lambda$.
2. By Lemma 5.1, the probability mass function V of the sum of independent Bernoulli variables with mean p_i , is c -log-concave with $c = (\sum_j p_j/(1-p_j))^{-1}$.
3. If $V(x) = \binom{n+x-1}{x} p^x (1-p)^n$ is negative binomial, then direct calculation gives

$$\mathcal{E}^{(V)}(x) = \frac{(n-1)}{p(n+x)(n+x-1)},$$

which tends to zero as $x \rightarrow \infty$. Hence $V(x)$ is only c -log-concave with $c = 0$.

One final remark is that no mass function with mean $\mathbb{E}V$ can be c -log-concave for $c > 1/(\mathbb{E}V)$. Hence the value $1/\lambda$ found for Π_λ in Example 5.2 is an extreme one.

Lemma 5.3. *If V is c -log-concave, then $c \leq 1/(\mathbb{E}V)$.*

Proof. Since $\mathcal{E}^{(V)}$ is a finite difference, we sum the collapsing sum to obtain

$$\frac{V(x)}{V(x+1)} = \sum_{y=0}^x \left(\frac{V(y)}{V(y+1)} - \frac{V(y-1)}{V(y)} \right) \geq (x+1)c,$$

by assumption. Rearranging and summing we obtain that

$$1 = \sum_{x=0}^{\infty} V(x) \geq \sum_{x=0}^{\infty} (x+1)V(x+1)c = c(\mathbb{E}V),$$

and the result follows. □

Note further that in some settings it may be natural to assume that $\mathcal{E}^{(V)}(x)$ is increasing in x . Direct substitution shows that this is equivalent to the property that

$$V(x)^2 V(x-1) - 2V(x-1)^2 V(x+1) + V(x+1)V(x)V(x-2) \geq 0, \quad \text{for all } x \geq 0. \quad (33)$$

In [17], this property (referred to there as “Property $C_1(k)$ ”) is shown by induction to hold when V is the probability mass function of the sum of independent Bernoulli variables, and it is natural to assume that (33) holds in a more general setting than this.

6. Proof of the log-Sobolev inequality, Theorem 1.3

Proof of Theorem 1.3. Given a fixed probability mass function $V(x)$ and a function f with $\sum_{x=0}^{\infty} V(x)f(x) = \mu_{V,f}$, we consider function f_t evolving as (21), that is with $f_0 \equiv f$ and

$$\frac{\partial}{\partial t} f_t(x) = L_V f_t(x).$$

Note that, by ergodicity, $\lim_{t \rightarrow \infty} f_t(x) = \sum_{x=0}^{\infty} V(x)f(x) = \mu_{V,f}$. We consider the function

$$\Theta(t) = \sum_{x=0}^{\infty} V(x) f_t(x) \log f_t(x), \quad (34)$$

and obtain that (as in [7]):

$$\begin{aligned} \Theta'(t) &= \sum_{x=0}^{\infty} V(x) L_V f_t(x) \log f_t(x) + \sum_{x=0}^{\infty} V(x) f_t(x) \frac{L_V f_t(x)}{f_t(x)} \\ &= - \sum_{x=0}^{\infty} V(x) (f_t(x+1) - f_t(x)) (\log f_t(x+1) - \log f_t(x)). \end{aligned} \quad (35)$$

This follows by cancellation, since $\sum_{x=0}^{\infty} V(x) L_V h(x) = 0$ for any h , and by taking $f = f_t$ and $g = \log f_t$ in (22). Since both terms in brackets in (35) have the same sign, we conclude that $\Theta'(t) = - \sum_{x=0}^{\infty} V(x) \Gamma_1^{(V)}(f_t, \log f_t)(x) \leq 0$ (this is the term arising in (10)). However the absence of a chain rule of the type (18) means that we cannot write it in a form where Condition 3 can be directly applied. However, we calculate a further derivative by hand.

In fact, we consider the derivative of a related term, which we think of as only part of the expression for $\Theta'(t)$. That is, we write

$$\psi(t) = \sum_{x=0}^{\infty} V(x) \left(f_t(x+1) \log \left(\frac{f_t(x+1)}{f_t(x)} \right) - f_t(x+1) + f_t(x) \right).$$

Using the fact that for functions g and h , $(g \log(g/h) - g)' = g' \log(g/h) - gh'/h$, by relabelling in the usual way we deduce that

$$\psi'(t) = \sum_{x=0}^{\infty} V(x) (L_V f_t(x+1) - L_V f_t(x)) \log \left(\frac{f_t(x+1)}{f_t(x)} \right) \quad (36)$$

$$+ \sum_{x=0}^{\infty} V(x) L_V f_t(x) \left(\log \left(\frac{f_t(x+1)}{f_t(x)} \right) - \frac{f_t(x+1)}{f_t(x)} \right). \quad (37)$$

By taking $g(x) = \log f_t(x)$ and $f(x) = f_t(x)$ in (31), we deal with (36), and by taking $f(x) = f_t(x)$ and

$$g(x) = \log \left(\frac{f_t(x+1)}{f_t(x)} \right) - \frac{f_t(x+1)}{f_t(x)}$$

in (22), we deal with (37). Adding the results of these manipulations together, we deduce that

$$\psi'(t) = - \sum_{x=0}^{\infty} V(x) \mathcal{E}^{(V)}(x) (f_t(x+1) - f_t(x)) \log \left(\frac{f_t(x+1)}{f_t(x)} \right) \quad (38)$$

$$+ \sum_{x=0}^{\infty} V(x) f_t(x+1) w \left(\frac{f_t(x)f_t(x+2)}{f_t(x+1)^2}; \frac{f_t(x)}{f_t(x+1)} \right), \quad (39)$$

where $w(U; s) = -(U/s - 1) \log U + (1 - U)(1 - 1/s)$. Lemma 6.1 gives that the term (39) is negative. (Note that this term is zero if $f_t(x) = \exp(ax + b)$, which contributes to the sharpness result discussed in Remark 1.4.) Further, by assumption, we bound (38) from above on replacing $\mathcal{E}^{(V)}(x)$ by c . In other words, we deduce by comparison with (35) that $\psi'(t) \leq c\Theta'(t)$, or that

$$(-\Theta'(t)) \leq \frac{1}{c}(-\psi'(t)).$$

We deduce

$$\text{Ent}_V(f) = \Theta(0) - \Theta(\infty) = \int_0^\infty -\Theta'(t) dt \leq \frac{1}{c} \int_0^\infty (-\psi'(t)) dt = \frac{1}{c} \psi(0)$$

and the result follows. \square

Lemma 6.1. *The function $w(U; s) = -(U/s - 1) \log U + (1 - U)(1 - 1/s) \leq 0$ for all $s, U \geq 0$, with equality if and only if $U = 1$.*

Proof. For fixed s , we observe that $w(1; s) = 0$, that

$$\frac{\partial}{\partial U} w(U; s) \Big|_{U=1} = 0$$

and $w(U; s)$ is a strictly concave function, since

$$\frac{\partial^2}{\partial U^2} w(U; s) = -\frac{s + U}{sU^2}. \quad \square$$

Remark 6.2. If V has support on a finite interval, a version of Theorem 1.3 should still hold, at least for a class of functions f . In brief, define $V_\epsilon := V \star \Pi_\epsilon$ to be the convolution of V with a Poisson mass function of mean ϵ . If V is c -log-concave, then for any given δ , the V_ϵ will be $(c - \delta)$ -log-concave for ϵ sufficiently small. Hence, we can apply Theorem 1.3 to V_ϵ (which is supported on the whole of \mathbb{Z}_+ as required) to obtain a bound on $\text{Ent}_{V_\epsilon}(f)$.

Further, by continuity arguments using dominated convergence $\text{Ent}_{V_\epsilon}(f)$ will converge to $\text{Ent}_V(f)$ for well-behaved f , and the resulting upper bound will also converge. However, we omit further discussion of this and the correct class of f to use for the sake of brevity.

7. Proof of the Poincaré inequality, Theorem 1.5

We show that the Poincaré inequality is equivalent to the integrated BE(c) condition, using a standard argument (see e.g. Proposition 4.8.3 of [3]).

Proof of Theorem 1.5. First, we assume the integrated BE(c) condition, and write $\Lambda(t) = \sum_{x=0}^\infty V(x) f_t(x)^2$. By Lemma 3.4 and (26)

$$\Lambda'(t) = 2 \sum_{x=0}^\infty V(x) f_t(x) L_V f_t(x) = -2 \sum_{x=0}^\infty V(x) (\Delta f_t)(x)^2 = -2 \sum_{x=0}^\infty V(x) \Gamma_1^{(V)}(f_t, f_t).$$

Similarly, since $\frac{\partial}{\partial t}$ commutes with L_V by the form of the Q -matrix in Definition 3.1,

$$\begin{aligned} \Lambda''(t) &= 2 \sum_{x=0}^\infty V(x) L_V f_t(x) L_V f_t(x) + 2 \sum_{x=0}^\infty V(x) f_t(x) L_V^2 f_t(x) \\ &= 4 \sum_{x=0}^\infty V(x) (L_V f_t(x))^2 = 4 \sum_{x=0}^\infty V(x) \Gamma_2^{(V)}(f_t, f_t). \end{aligned}$$

The integrated $\text{BE}(c)$ condition applied to the function f_t tells us that $\Lambda''(t) \geq -2c\Lambda'(t)$. This tells us that

$$\text{Var}_V(f) = \Lambda(0) - \Lambda(\infty) = \int_0^\infty -\Lambda'(t) dt \leq \frac{1}{2c} \int_0^\infty \Lambda''(t) dt = \frac{1}{2c} (-\Lambda'(0)) = \frac{1}{c} \sum_{x=0}^\infty V(x) (\Delta f)(x)^2,$$

and the result follows.

Second, if the Poincaré inequality holds, we deduce the integrated $\text{BE}(c)$ condition, since without loss of generality we can consider for any f with $\sum_x V(x) f(x) = 0$, for which

$$\sum_{x=0}^\infty V(x) \Gamma_1^{(V)}(f, f)(x) = - \sum_{x=0}^\infty V(x) f(x) L_V f(x) \quad (40)$$

$$\leq \sqrt{\sum_{x=0}^\infty V(x) f(x)^2} \sqrt{\sum_{x=0}^\infty V(x) L_V f(x)^2} \quad (41)$$

$$\leq \sqrt{\frac{1}{c} \sum_{x=0}^\infty V(x) \Gamma_1^{(V)}(f, f)(x)} \sqrt{\sum_{x=0}^\infty V(x) \Gamma_2^{(V)}(f, f)(x)}, \quad (42)$$

where (40) follows by (29), (41) follows by Cauchy–Schwarz, and (42) follows since by (12) $\sum_{x=0}^\infty V(x) f(x)^2 = \text{Var}_V(f) \leq \frac{1}{c} \sum_{x=0}^\infty V(x) \Delta f(x)^2 = \frac{1}{c} \sum_{x=0}^\infty V(x) \Gamma_1^{(V)}(f, f)(x)$. \square

Theorem 1.5 shows that if V satisfies the integrated $\text{BE}(c)$ condition, then the Poincaré constant of V is $\leq 1/c$. In comparison [12, Corollary 2.4], which was proved using arguments based on stochastic ordering and size-biasing, shows that if V is ULC then the Poincaré constant of V is less than or equal to $\mathbb{E}V$. Lemma 5.1 and Proposition 4.2 show ULC implies the integrated $\text{BE}(c)$ condition, hence the assumptions of the present paper are weaker than in [12]. However, Lemma 5.3 shows that $\mathbb{E}V \leq 1/c$, so here we prove a weaker bound on the Poincaré constant. It would be of interest to know if the two approaches can be synthesised, or if the results are each optimal under their own assumptions.

8. Consequences of Theorem 1.3

We briefly discuss some results which follow from Theorem 1.3, including a concentration of measure inequality, decay of entropy and a form of hypercontractivity.

8.1. Concentration of measure

We prove a concentration of measure result by adapting the argument used to prove [4, Proposition 10], and deduce the following bound:

Proposition 8.1. *Fix probability mass function V , and suppose that for all functions f with positive values, Equation (4) holds, that is:*

$$\text{Ent}_V(f) \leq \frac{1}{c} \sum_{x=0}^\infty V(x) f(x+1) \left(\log \left(\frac{f(x+1)}{f(x)} \right) - 1 + \frac{f(x)}{f(x+1)} \right). \quad (43)$$

Then, writing $h(s) = (1+s) \log(1+s) - s$, for any function g with $\sup_x |g(x+1) - g(x)| \leq 1$:

$$V(\{g \geq \mathbb{E}_V g + t\}) \leq \exp \left(-\frac{h(ct)}{c} \right). \quad (44)$$

Proof. Define the function $G(\tau) = \sum_{x=0}^{\infty} V(x)e^{\tau g(x)}$, and the related function $H(\tau) = (\log G(\tau))/\tau$. Taking $f(x) = e^{\tau g(x)}$ in (43) we deduce that:

$$\begin{aligned}
 \tau^2 G(\tau) H'(\tau) &= \tau G'(\tau) - G(\tau) \log G(\tau) \\
 &= \sum_{x=0}^{\infty} V(x) \tau g(x) e^{\tau g(x)} - G(\tau) \log G(\tau) \\
 &= \text{Ent}_V(e^{\tau g}) \\
 &\leq \frac{1}{c} \sum_{x=0}^{\infty} V(x) [\tau e^{\tau g(x+1)} (g(x+1) - g(x)) - e^{\tau g(x+1)} + e^{\tau g(x)}] \\
 &= \frac{1}{c} \sum_{x=0}^{\infty} V(x) e^{\tau g(x)} \varphi(\tau \Delta g(x)),
 \end{aligned} \tag{45}$$

where $\varphi(u) = ue^u - e^u + 1 \geq 0$ and $\Delta g(x) = g(x+1) - g(x)$.

Since $\varphi'(u) = ue^u$, which has the same sign as u , we know that taking $\tau \geq 0$ and for $v \in (-\tau, \tau)$, the $\varphi(v) \leq \max(\varphi(\tau), \varphi(-\tau)) = \varphi(\tau)$, where this last inequality follows since $\varphi(v) - \varphi(-v)$ is increasing on $v \geq 0$, and hence is ≥ 0 . Using this, we can rewrite (45) in the form $\tau^2 G(\tau) H'(\tau) \leq \frac{\varphi(\tau)}{c} G(\tau)$, which we can integrate to deduce that for any $\sigma \geq 0$:

$$H(\sigma) - H(0) = \int_0^{\sigma} H'(\tau) d\tau \leq \frac{1}{c} \int_0^{\sigma} \frac{\varphi(\tau)}{\tau^2} d\tau = \frac{1}{c} \frac{e^{\sigma} - \sigma - 1}{\sigma}. \tag{46}$$

This can be rearranged to give an upper bound on $G(\sigma)$. As in [4, Proposition 10], we can use a standard Chernoff bounding argument, based on the fact that $H(0) = \mathbb{E}_V g$ and using Markov's inequality to deduce that for any $\sigma > 0$:

$$V(\{g \geq \mathbb{E}_V g + t\}) \leq \frac{\mathbb{E}_V e^{\sigma g}}{e^{\sigma(\mathbb{E}_V g + t)}} = \frac{G(\sigma)}{e^{\sigma(\mathbb{E}_V g + t)}} \leq \exp\left(\frac{e^{\sigma} - \sigma - 1}{c} - \sigma t\right).$$

We make the optimal choice of σ here, that is $\sigma = \log(1 + ct)$, to deduce the result. \square

Note this function h commonly occurs in concentration of measure results in different settings, including Bennett's inequality (see e.g. [29, Theorem 9]), work of Houdré and co-authors based on the "covariance method" (see e.g. [19, Eq. (1.6)]) and recent work on discrete random variables using a tail condition under coupling [11, Theorem 3.3].

Remark 8.2. Proposition 8.1 shows that Theorem 1.3 can provide practical improvements to results of the form (9). To be specific, [4, Proposition 10] shows that if (9) holds, then, under the same condition on Δg :

$$V(\{g \geq \mathbb{E}_V g + t\}) \leq \exp\left(-\frac{k(ct)}{c}\right), \tag{47}$$

where $k(u) = u \log(1 + u)/4$. Proposition 8.1 therefore strengthens (47) under the c -log-concavity condition, Condition 1, since $h(u) \geq 2k(u)$ for all u . This strengthening comes from the fact that the expression of (46) is significantly smaller than the bound of $\frac{1}{2c}(e^{2\sigma} - 1)$ which follows by the argument of [4, Proposition 10]. Note that (46) is sharp, in the sense that equality holds when taking $V = \Pi_{\lambda}$ and $g(x) = x$, as follows from the sharpness of Theorem 1.3 discussed in Remark 1.4, item 6.

8.2. Decay of entropy and hypercontractivity

We briefly discuss how the log-Sobolev inequality, Theorem 1.3, implies further results for related processes, in a standard way. Motivated by the paper [33], which considered pure thinning, we consider probability measures

evolving as the “death” part of the birth and death process. That is, for fixed V , we consider probability distributions such that:

$$\frac{\partial}{\partial t} V_t(x) = \alpha_t (V_t(x) - V_t(x-1)), \quad (48)$$

$$\frac{\partial}{\partial t} p_t(x) = \alpha_t \left(\frac{V_t(x)}{V_t(x+1)} p_t(x+1) - \frac{V_t(x-1)}{V_t(x)} p_t(x) \right). \quad (49)$$

Proposition 8.3. *If $V_t(x)$ satisfies $\mathcal{E}^{(V_t)}(x) \geq c_t$ for all x then*

$$D(p_t \parallel V_t) \leq D(p \parallel V) \exp\left(-\int_0^t \alpha_s c_s ds\right).$$

Proof. Writing $K_t = (\sum_{x=0}^{\infty} p_t(x+1) V_t(x) / V_t(x+1))^{-1}$ and $p_t^\#(x) = K_t p_t(x+1) V_t(x) / V_t(x+1)$ relabelling gives:

$$\begin{aligned} & \frac{\partial}{\partial t} D(p_t \parallel V_t) \\ &= \alpha_t \sum_{x=0}^{\infty} \left(\frac{V_t(x)}{V_t(x+1)} p_t(x+1) - \frac{V_t(x-1)}{V_t(x)} p_t(x) \right) \log\left(\frac{p_t(x)}{V_t(x)}\right) - \frac{V_t(x) - V_t(x-1)}{V_t(x)} p_t(x) \\ &= \alpha_t \left(\sum_{x=0}^{\infty} \frac{V_t(x) p_t(x+1)}{V_t(x+1)} \log\left(\frac{p_t(x) V_t(x+1)}{V_t(x) p_t(x+1)}\right) - 1 + \frac{1}{K_t} \right) \\ &= -\frac{\alpha_t}{K_t} \left(D(p_t^\# \parallel p_t) + \log \frac{1}{K_t} - 1 + K_t \right) \\ &\leq -\alpha_t c_t D(p_t \parallel V_t), \end{aligned}$$

where the last inequality follows using the form of the log-Sobolev inequality given by Equation (6). \square

Example 8.4. Taking $V_t = \Pi_{\lambda(t)}$, a Poisson mass function with mean $\lambda(t) = \lambda e^{-t}$, then (48) holds with $\alpha_t = \lambda(t)$, and we know that $c_t = 1/\lambda(t)$. Hence, P_t becomes the mass function P thinned by e^{-t} (see [12] for a discussion of this operation), and we can deduce that

$$D(p_t \parallel V_t) \leq D(p \parallel V) e^{-t}. \quad (50)$$

We also illustrate Theorem 1.3 by using it to prove a form of hypercontractivity, using a standard argument (see e.g. [5, Theorem 11] and [3, page 246]).

Proposition 8.5. *Consider a sequence of probability measures evolving as in (48) and a sequence of functions evolving in a related way:*

$$\begin{aligned} & \frac{\partial}{\partial t} V_t(x) = \alpha_t (V_t(x) - V_t(x-1)), \\ & \frac{\partial}{\partial t} g_t(x) = \alpha_t \frac{V_t(x-1)}{V_t(x)} (g_t(x) - g_t(x-1)). \end{aligned} \quad (51)$$

If V_t satisfies the new modified log-Sobolev inequality, Equation (4) with constant c_t then writing $q(t) = p \exp(-\int_0^t \alpha_s c_s ds)$ and $\|f\|_{U,p} = (\sum_{x=0}^{\infty} U(x) f(x)^p)^{1/p}$ then

$$\|\exp(g)\|_{V,p} \leq \|\exp(g_t)\|_{V_t,q(t)}. \quad (52)$$

(Note that $q(t) \leq p$.)

Proof. As in [3,5], we consider the functional $\Lambda(q, t) := \sum_{x=0}^{\infty} V_t(x) \exp(qg_t(x))$. The key is to express

$$\begin{aligned} \frac{\partial}{\partial q} \Lambda(q, t) &= \sum_{x=0}^{\infty} V_t(x) \exp(qg_t(x)) g_t(x) \\ &= \frac{1}{q} \text{Ent}_{V_t}(\exp(qg_t)) + \frac{1}{q} \Lambda(q, t) \log \Lambda(q, t) \end{aligned}$$

and using (48) and (51) to recognise that $\frac{\partial}{\partial t} \Lambda(q, t)$ equals

$$\alpha_t \left[\sum_{x=0}^{\infty} V_t(x+1) (q \exp(qg_t(x+1)) (g_t(x+1) - g_t(x)) - \exp(qg_t(x+1)) + \exp(qg_t(x))) \right].$$

Taking $f = \exp(qg_t)$ in (4) we deduce that

$$-\partial_q \log \Lambda(q, t) + \frac{1}{\alpha_t c_t q} \partial_t \log \Lambda(q, t) + \frac{1}{q} \log \Lambda(q, t) \geq 0. \quad (53)$$

Using this, we can consider the behaviour of $u(t) := \log \Lambda(q(t), t)/q(t)$. Taking a derivative with respect to t , using the fact that $q'(t)/q(t) = -c_t \alpha_t$, we obtain that

$$\begin{aligned} u'(t) &= \frac{q'(t)}{q(t)} \partial_q \log \Lambda(q(t), t) + \frac{1}{q(t)} \partial_t \log \Lambda(q(t), t) - \frac{q'(t)}{q(t)^2} \log \Lambda(q(t), t) \\ &= c_t \alpha_t \left(-\partial_q \log \Lambda(q(t), t) + \frac{1}{q(t) c_t \alpha_t} \partial_t \log \Lambda(q(t), t) + \frac{1}{q(t)} \log \Lambda(q(t), t) \right) \\ &\geq 0, \end{aligned}$$

where the final inequality follows from (53). Since $\exp(u(t)) = \|\exp(g_t)\|_{V_t, q(t)}$, we deduce that the q norm is increasing as required. \square

Note that the definition of $q(t)$ involves the same exponential expression as Proposition 8.3.

Example 8.6. As in Example 8.4, we can consider $V_t = \Pi_{\lambda(t)}$, a Poisson mass function with mean $\lambda(t) = \lambda e^{-t}$, and notice that (51) is satisfied by the Poisson–Charlier polynomials $c_k(x; \lambda(t))$. Further, since $c_t \alpha_t = 1$, we take $q(t) = p \exp(-t)$.

We deduce that Proposition 8.5 is sharp, taking $g_t = (x - \lambda(t))/\lambda(t)$ to be the Poisson–Charlier polynomial of degree 1. In this case

$$\begin{aligned} \Lambda(q(t), t) &= \sum_{x=0}^{\infty} \frac{\exp(-\lambda(t)) \lambda(t)^x}{x!} \exp\left(\frac{q(t)(x - \lambda(t))}{\lambda(t)}\right) \\ &= \exp(-q(t) - \lambda(t) + \lambda(t) e^{q(t)/\lambda(t)}) = \exp(-q(t)C), \end{aligned}$$

where $C = 1 + \lambda/p - \lambda/pe^{p/\lambda}$, using the fact that $\lambda(t)/q(t) \equiv \lambda/p$. Hence the $q(t)$ norm is constant, and Proposition 8.5 is sharp. (This sharpness corresponds to the sharpness of the new modified log-Sobolev inequality for functions of the form $f(x) = \exp(ax + b)$, as discussed in Remark 1.4, item 6.)

9. Extension to random variables on \mathbb{Z}_+^d

It would be of considerable interest to extend this work to the more general setting of probability measures on graphs, where curvature and related issues are topics of active research. For example, [1] uses the Bakry–Émery Γ -calculus to

deduce log-Sobolev inequalities on the discrete cube, \mathbb{Z}^d and general graphs of uniformly bounded degree. The paper [23] defines curvature for discrete graphs, and shows that controlling this curvature allows results including Poincaré and log-Sobolev inequalities to be deduced. Lin and Yau [26] compare the two forms of curvature discussed by Joulin [21], in the context of graphs.

We briefly describe how the c -log-concavity condition, Condition 1, extends to the setting of probability measures on \mathbb{Z}_+^d . We deduce an integrated Bakry–Émery condition, and hence a Poincaré inequality, and explain the issues with proving a modified log-Sobolev inequality in the form of Theorem 1.3.

Fix a reference measure $V(\mathbf{x})$ which is positive for all $\mathbf{x} \in \mathbb{Z}_+^d$, and write \mathbf{e}_i for the i th unit vector. Further, for all $1 \leq i, j \leq d$ and for a given function f we define

$$\mathcal{E}_{ij}^{(V)}(\mathbf{x}) = \frac{V(\mathbf{x} + \mathbf{e}_j - \mathbf{e}_i)}{V(\mathbf{x} + \mathbf{e}_j)} - \frac{V(\mathbf{x} - \mathbf{e}_i)}{V(\mathbf{x})}, \quad (54)$$

$$L_{ij}f(\mathbf{x}) = f(\mathbf{x} + \mathbf{e}_j) - f(\mathbf{x} + \mathbf{e}_j - \mathbf{e}_i) - f(\mathbf{x}) + f(\mathbf{x} - \mathbf{e}_i). \quad (55)$$

Notice that these quantities are not symmetric in i and j , however for each \mathbf{y} and each c we can define a symmetric matrix by

$$\begin{aligned} E_{ij}^{(V,c)}(\mathbf{y}) &:= \frac{V(\mathbf{y} - \mathbf{e}_i)V(\mathbf{y} - \mathbf{e}_j)}{V(\mathbf{y})} - V(\mathbf{y} - \mathbf{e}_i - \mathbf{e}_j) - c\mathbb{I}(i = j)V(\mathbf{y} - \mathbf{e}_j) \\ &= V(\mathbf{y} - \mathbf{e}_j)(\mathcal{E}_{ij}^{(V)}(\mathbf{y} - \mathbf{e}_j) - c\mathbb{I}(i = j)) \end{aligned} \quad (56)$$

writing $\mathbb{I}(i = j)$ for the entries of the identity matrix. Consider a process which (for all \mathbf{x} and all i) jumps from \mathbf{x} to $\mathbf{x} + \mathbf{e}_i$ at rate 1 and from \mathbf{x} to $\mathbf{x} - \mathbf{e}_i$ at rate $V(\mathbf{x} - \mathbf{e}_i)/V(\mathbf{x})$ (where by convention $V(\mathbf{y}) = 0$ if any component of \mathbf{y} is -1). This corresponds to defining

$$L_V f(\mathbf{x}) = \sum_{i=1}^d (f(\mathbf{x} + \mathbf{e}_i) - f(\mathbf{x})) - \frac{V(\mathbf{x} - \mathbf{e}_i)}{V(\mathbf{x})} (f(\mathbf{x}) - f(\mathbf{x} - \mathbf{e}_i)). \quad (57)$$

The key is to observe that an analogue of (28) holds, that is direct calculation gives that for any j :

$$\begin{aligned} L_V f(\mathbf{x} + \mathbf{e}_j) - L_V f(\mathbf{x}) \\ = \sum_{i=1}^d \left(L_{ij}f(\mathbf{x} + \mathbf{e}_i) - L_{ij}f(\mathbf{x}) \frac{V(\mathbf{x} - \mathbf{e}_i)}{V(\mathbf{x})} - \mathcal{E}_{ij}^{(V)}(\mathbf{x})(f(\mathbf{x} + \mathbf{e}_j) - f(\mathbf{x} + \mathbf{e}_j - \mathbf{e}_i)) \right). \end{aligned} \quad (58)$$

We deduce that:

Proposition 9.1. *If for some c , the matrix $E^{(V,c)}(\mathbf{y})$ of Equation (56) is positive definite for all \mathbf{y} , then for any function f :*

$$\sum_{\mathbf{x} \in \mathbb{Z}_+^d} V(\mathbf{x}) \Gamma_2(f, f)(\mathbf{x}) \geq c \sum_{\mathbf{x} \in \mathbb{Z}_+^d} V(\mathbf{x}) \Gamma_1(f, f)(\mathbf{x}). \quad (59)$$

Proof. First we observe that (by relabelling)

$$\begin{aligned} \sum_{\mathbf{x} \in \mathbb{Z}_+^d} V(\mathbf{x}) \Gamma_1(f, g)(\mathbf{x}) &= \sum_{\mathbf{x} \in \mathbb{Z}_+^d} V(\mathbf{x}) \sum_{j=1}^d (f(\mathbf{x} + \mathbf{e}_j) - f(\mathbf{x}))(g(\mathbf{x} + \mathbf{e}_j) - g(\mathbf{x})) \\ &= \sum_{j=1}^d \sum_{\mathbf{y} \in \mathbb{Z}_+^d} V(\mathbf{y} - \mathbf{e}_j)(f(\mathbf{y}) - f(\mathbf{y} - \mathbf{e}_j))(g(\mathbf{y}) - g(\mathbf{y} - \mathbf{e}_j)). \end{aligned} \quad (60)$$

Using this, we can deduce a d -dimensional version of (32), namely

$$\begin{aligned}
& \sum_{\mathbf{x} \in \mathbb{Z}_+^d} V(\mathbf{x}) \Gamma_2(f, g)(\mathbf{x}) \\
&= - \sum_{\mathbf{x} \in \mathbb{Z}_+^d} V(\mathbf{x}) \sum_{j=1}^d (L_V f(\mathbf{x} + \mathbf{e}_j) - L_V f(\mathbf{x})) (g(\mathbf{x} + \mathbf{e}_j) - g(\mathbf{x})) \\
&= \sum_{\mathbf{x} \in \mathbb{Z}_+^d} V(\mathbf{x}) \sum_{i,j} L_{ij} f(\mathbf{x} + \mathbf{e}_i) L_{ij} g(\mathbf{x} + \mathbf{e}_i) \\
&\quad + \sum_{\mathbf{x} \in \mathbb{Z}_+^d} V(\mathbf{x}) \sum_{i=1}^d \sum_{j=1}^d \mathcal{E}_{ij}^{(V)}(\mathbf{x}) (f(\mathbf{x} + \mathbf{e}_j) - f(\mathbf{x} + \mathbf{e}_j - \mathbf{e}_i)) (g(\mathbf{x} + \mathbf{e}_j) - g(\mathbf{x})) \\
&\geq \sum_{i=1}^d \sum_{j=1}^d \sum_{\mathbf{y} \in \mathbb{Z}_+^d} V(\mathbf{y} - \mathbf{e}_j) \mathcal{E}_{ij}^{(V)}(\mathbf{y} - \mathbf{e}_j) (f(\mathbf{y}) - f(\mathbf{y} - \mathbf{e}_i)) (g(\mathbf{y}) - g(\mathbf{y} - \mathbf{e}_j)), \tag{61}
\end{aligned}$$

so subtracting c times (60) from (61) and taking $f = g$ we obtain that

$$\begin{aligned}
& \sum_{\mathbf{x} \in \mathbb{Z}_+^d} V(\mathbf{x}) \Gamma_2(f, f)(\mathbf{x}) - c \sum_{\mathbf{x} \in \mathbb{Z}_+^d} V(\mathbf{x}) \Gamma_1(f, f)(\mathbf{x}) \\
&\geq \sum_{\mathbf{y} \in \mathbb{Z}_+^d} \left[\sum_{i=1}^d \sum_{j=1}^d E_{ij}^{(V,c)}(\mathbf{y}) (f(\mathbf{y}) - f(\mathbf{y} - \mathbf{e}_j)) (f(\mathbf{y}) - f(\mathbf{y} - \mathbf{e}_i)) \right],
\end{aligned}$$

and the term in square brackets is positive for each \mathbf{y} , by positive-definiteness. \square

Remark 9.2. If $V(\mathbf{x}) = \prod_{k=1}^d V_k(x_k)$ is formed as the product of independent measures in each coordinate, then $E_{ij}^{(V,c)}(\mathbf{y}) \equiv 0$ for $i \neq j$. Further, if each V_i is c -log-concave then each entry $E_{ii}^{(V,c)}(\mathbf{y}) = V_i(y_i - 1)(\mathcal{E}^{(V_i)}(y_i - 1) - c) \geq 0$, so the condition of Proposition 9.1 is satisfied. This mirrors the tensorization result of [13, Theorem 1.3], which was used to prove a sharp bound on the Ricci curvature for the hypercube $\{0, 1\}^d$.

Hence, repeating the proof of Theorem 1.5, we can deduce that the positive definiteness of $E^{(V,c)}(\mathbf{y})$ for all \mathbf{y} is enough to imply that a d -dimensional Poincaré inequality holds with constant $\leq 1/c$.

A more detailed argument shows that many of the arguments used in Section 6 to prove the new modified log-Sobolev inequality Theorem 1.3 carry over. That is, we consider functions $f_t(\mathbf{x})$ evolving as $\frac{\partial}{\partial t} f_t(\mathbf{x}) = L_V f_t(\mathbf{x})$, for L_V as defined in (57). Again, taking $\Theta(t) := \sum_{\mathbf{x} \in \mathbb{Z}_+^d} V(\mathbf{x}) f_t(\mathbf{x}) \log f_t(\mathbf{x})$, we obtain that

$$\Theta'(t) = - \sum_{\mathbf{x} \in \mathbb{Z}_+^d} V(\mathbf{x}) \sum_{j=1}^d (f_t(\mathbf{x} + \mathbf{e}_j) - f_t(\mathbf{x})) (\log f_t(\mathbf{x} + \mathbf{e}_j) - \log f_t(\mathbf{x})). \tag{62}$$

Similarly, writing

$$\psi(t) := \sum_{\mathbf{x} \in \mathbb{Z}_+^d} V(\mathbf{x}) \sum_{j=1}^d \left(f_t(\mathbf{x} + \mathbf{e}_j) \log \left(\frac{f_t(\mathbf{x} + \mathbf{e}_j)}{f_t(\mathbf{x})} \right) - f_t(\mathbf{x} + \mathbf{e}_j) + f_t(\mathbf{x}) \right), \tag{63}$$

an involved analysis using the expressions above shows that

$$\begin{aligned}\psi'(t) &= \sum_{\mathbf{x} \in \mathbb{Z}_+^d} V(\mathbf{x}) \sum_{i,j=1}^d f(\mathbf{x} + \mathbf{e}_j) w(U_{ij}(\mathbf{x}), s_i(\mathbf{x})) \\ &\quad - \sum_{\mathbf{x} \in \mathbb{Z}_+^d} V(\mathbf{x}) \sum_{i,j=1}^d \mathcal{E}_{ij}^{(V)}(\mathbf{x}) (f_t(\mathbf{x} + \mathbf{e}_j) - f_t(\mathbf{x} + \mathbf{e}_j - \mathbf{e}_i)) (\log f_t(\mathbf{x} + \mathbf{e}_j) - \log f_t(\mathbf{x})),\end{aligned}$$

where as before $w(U; s) = -(U/s - 1) \log U + (1 - U)(1 - 1/s) \geq 0$ and we write $U_{ij}(\mathbf{x}) = f(\mathbf{x})f(\mathbf{x} + \mathbf{e}_i + \mathbf{e}_j)/(f(\mathbf{x} + \mathbf{e}_i)f(\mathbf{x} + \mathbf{e}_j))$ and $s_i(\mathbf{x}) = f(\mathbf{x})/f(\mathbf{x} + \mathbf{e}_i)$. We deduce that

$$c\Theta'(t) - \psi'(t) = \sum_{\mathbf{y} \in \mathbb{Z}_+^d} \left[\sum_{i=1}^d \sum_{j=1}^d E_{ij}^{(V,c)}(\mathbf{y}) (f_t(\mathbf{y}) - f_t(\mathbf{y} - \mathbf{e}_j)) (\log f_t(\mathbf{y}) - \log f_t(\mathbf{y} - \mathbf{e}_i)) \right]. \quad (64)$$

Unfortunately, positive definiteness of $E^{(V,c)}$ is not sufficient to guarantee the positivity of (64) required to deduce the log-Sobolev inequality. If (as in Remark 9.2) V is the product of c -log-concave mass functions, then $E^{(V,c)}$ becomes diagonal with positive entries. The positivity of (64) follows from the fact that \log is a monotone function, meaning that $(f_t(\mathbf{y}) - f_t(\mathbf{y} - \mathbf{e}_j))$ and $(\log f_t(\mathbf{y}) - \log f_t(\mathbf{y} - \mathbf{e}_j))$ have the same sign.

It remains an interesting problem to characterize probability mass functions on \mathbb{Z}_+^d (and indeed for general graph settings) for which some form of Theorem 1.3 holds.

Acknowledgements

The author thanks the University of Bristol for funding to attend the conference “When Dominique Bakry is 60” at Université Paul Sabatier Toulouse in December 2014. He also thanks the Institute for Mathematics and Its Applications for the invitation and funding to speak at the workshop “Information Theory and Concentration Phenomena” in Minneapolis in April 2015. Attending talks and having discussions with organisers and fellow participants at these meetings greatly helped in understanding the topics discussed here. Many calculations in this paper were first performed using Mathematica. The author thanks the anonymous referees of this paper for making numerous extremely helpful suggestions.

References

- [1] C. Ané and M. Ledoux. On logarithmic Sobolev inequalities for continuous time random walks on graphs. *Probab. Theory Related Fields* **116** (4) (2000) 573–602. [MR1757600](#)
- [2] D. Bakry and M. Émery. Diffusions hypercontractives. In *Séminaire de Probabilités XIX* 177–206. *Lecture Notes in Math.* **1123**. Springer, Berlin, 1985. [MR0889476](#)
- [3] D. Bakry, I. Gentil and M. Ledoux. *Analysis and Geometry of Markov Diffusion Operators*. *Grundlehren der Mathematischen Wissenschaften* **348**. Springer, Cham, 2014. [MR3155209](#)
- [4] S. G. Bobkov and M. Ledoux. On modified logarithmic Sobolev inequalities for Bernoulli and Poisson measures. *J. Funct. Anal.* **156** (2) (1998) 347–365. [MR1636948](#)
- [5] S. G. Bobkov and P. Tetali. Modified logarithmic Sobolev inequalities in discrete settings. *J. Theoret. Probab.* **19** (2) (2006) 289–336. [MR2283379](#)
- [6] A.-S. Boudou, P. Caputo, P. Dai Pra and G. Posta. Spectral gap estimates for interacting particle systems via a Bochner-type identity. *J. Funct. Anal.* **232** (1) (2006) 222–258. [MR2200172](#)
- [7] P. Caputo, P. Dai Pra and G. Posta. Convex entropy decay via the Bochner–Bakry–Emery approach. *Ann. Inst. Henri Poincaré Probab. Stat.* **45** (3) (2009) 734–753. [MR2548501](#)
- [8] D. Chafaï. Binomial-Poisson entropic inequalities and the $M/M/\infty$ queue. *ESAIM Probab. Stat.* **10** (2006) 317–339. [MR2247924](#)
- [9] D. Chafaï and A. Joulin. Intertwining and commutation relations for birth–death processes. *Bernoulli* **19** (5A) (2013) 1855–1879. [MR3129037](#)
- [10] M. Chen. Estimation of spectral gap for Markov chains. *Acta Math. Sin. (N.S.)* **12** (4) (1996) 337–360. [MR1457859](#)

- [11] N. A. Cook, L. Goldstein and T. Johnson. Size biased couplings and the spectral gap for random regular graphs. Preprint, 2015. Available at [arXiv:1510.06013](https://arxiv.org/abs/1510.06013).
- [12] F. Daly and O. T. Johnson. Bounds on the Poincaré constant under negative dependence. *Statist. Probab. Lett.* **83** (2013) 511–518. [MR3006983](#)
- [13] M. Erbar and J. Maas. Ricci curvature of finite Markov chains via convexity of the entropy. *Arch. Ration. Mech. Anal.* **206** (2012) 997–1038. [MR2989449](#)
- [14] M. Fathi and J. Maas. Entropic Ricci curvature bounds for discrete interacting systems. *Ann. Appl. Probab.* **26** (2016) 1774–1806. [MR3513606](#)
- [15] L. Gross. Logarithmic Sobolev inequalities. *Amer. J. Math.* **97** (4) (1975) 1061–1083. [MR0420249](#)
- [16] A. Guionnet and B. Zegarlinski. Lectures on logarithmic Sobolev inequalities. In *Séminaire de Probabilités XXXVI* 1–134. *Lecture Notes in Math.* **1801**. Springer, Berlin, 2003. [MR1971582](#)
- [17] E. Hillion and O. T. Johnson. Discrete versions of the transport equation and the Shepp–Olkin conjecture. *Ann. Probab.* **44** (1) (2016) 276–306. [MR3456338](#)
- [18] E. Hillion, O. T. Johnson and Y. Yu. A natural derivative on $[0, n]$ and a binomial Poincaré inequality. *ESAIM Probab. Stat.* **16** (2014) 703–712. [MR3334010](#)
- [19] C. Houdré. Remarks on deviation inequalities for functions of infinitely divisible random vectors. *Ann. Probab.* **33** (3) (2002) 1223–1237. [MR1920106](#)
- [20] O. T. Johnson. Log-concavity and the maximum entropy property of the Poisson distribution. *Stochastic Process. Appl.* **117** (6) (2007) 791–802. [MR2327839](#)
- [21] A. Joulin. Poisson-type deviation inequalities for curved continuous-time Markov chains. *Bernoulli* **13** (2007) 782–798. [MR2348750](#)
- [22] C. Klaassen. On an inequality of Chernoff. *Ann. Probab.* **13** (3) (1985) 966–974. [MR0799431](#)
- [23] B. Klartag, G. Kozma, P. Ralli and P. Tetali. Discrete curvature and abelian groups. Preprint, 2015. Available at [arXiv:1501.00516](https://arxiv.org/abs/1501.00516).
- [24] I. Kontoyiannis, P. Harremoës and O. T. Johnson. Entropy and the law of small numbers. *IEEE Trans. Inform. Theory* **51** (2) (2005) 466–472. [MR2236061](#)
- [25] T. M. Liggett. Ultra logconcave sequences and negative dependence. *J. Combin. Theory Ser. A* **79** (2) (1997) 315–325. [MR1462561](#)
- [26] Y. Lin and S.-T. Yau. Ricci curvature and eigenvalue estimate on locally finite graphs. *Math. Res. Lett.* **17** (2) (2010) 343–356. [MR2644381](#)
- [27] A. Mielke. Geodesic convexity of the relative entropy in reversible Markov chains. *Calc. Var. Partial Differential Equations* **48** (1–2) (2013) 1–31. [MR3090532](#)
- [28] R. Pemantle. Towards a theory of negative dependence. *J. Math. Phys.* **41** (3) (2000) 1371–1390. [MR1757964](#)
- [29] M. Raginsky and I. Sason. Concentration of measure inequalities in information theory, communications and coding. *Found. Trends Commun. Inf. Theory* **10** (1–2) (2013) 1–246.
- [30] A. J. Stam. Some inequalities satisfied by the quantities of information of Fisher and Shannon. *Inf. Control* **2** (1959) 101–112. [MR0109101](#)
- [31] D. W. Walkup. Pólya sequences, binomial convolution and the union of random sets. *J. Appl. Probab.* **13** (1) (1976) 76–85. [MR0494391](#)
- [32] L. Wu. A new modified logarithmic Sobolev inequality for Poisson point processes and several applications. *Probab. Theory Related Fields* **118** (3) (2000) 427–438. [MR1800540](#)
- [33] Y. Yu. Monotonic convergence in an information-theoretic law of small numbers. *IEEE Trans. Inform. Theory* **55** (12) (2009) 5412–5422. [MR2597172](#)